

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/511, 244  
Source: PCI  
Date Processed by STIC: 02/02/2006

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 02/02/2006

PATENT APPLICATION: US/10/511,244

TIME: 12:34:04

Input Set : A:\UCSD1420-1.ST25.txt

Output Set: N:\CRF4\02022006\J511244.raw

3 <110> APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
 4 THE UNIVERSITY OF BRITISH COLUMBIA  
 5 SAREEN, Dipti  
 6 NEWTON, Gerald L.  
 7 FAHEY, Robert C.  
 8 BUCHMEIER, Nancy  
 9 STEFFEK, Micah  
 10 AV-GAY, Yossef  
 11 RAWAT, Mamta  
 12 KOLEDIN, Teresa  
 14 <120> TITLE OF INVENTION: METHODS OF USE OF THE ENZYMES OF MYCOTHIOIOL SYNTHESIS  
 16 <130> FILE REFERENCE: UCSD1420-1  
 18 <140> CURRENT APPLICATION NUMBER: US 10/511,244  
 C--> 19 <141> CURRENT FILING DATE: 2004-10-13  
 21 <150> PRIOR APPLICATION NUMBER: PCT/US 03/11539  
 22 <151> PRIOR FILING DATE: 2003-04-15  
 24 <150> PRIOR APPLICATION NUMBER: US 60/373,890  
 25 <151> PRIOR FILING DATE: 2002-04-19  
 27 <150> PRIOR APPLICATION NUMBER: US 60/373,079  
 28 <151> PRIOR FILING DATE: 2002-04-15  
 30 <160> NUMBER OF SEQ ID NOS: 49  
 32 <170> SOFTWARE: PatentIn version 3.1  
 34 <210> SEQ ID NO: 1  
 35 <211> LENGTH: 1236  
 36 <212> TYPE: DNA  
 37 <213> ORGANISM: Mycobacterium smegmatis  
 39 <400> SEQUENCE: 1  
 40 atgcaatcgt ggtcggcacc ggcgattccg gtgggtccgg gacgtggccc tgcgtgcgc 60  
 42 ctcttcgaca gcgctgatcg ccaggtccgg cccgtcacac cgggaccgac cgcaaccatg 120  
 44 tacgtgtgcg gcatcaccac atacgacgcg acccatctgg gtcacgccgc gacctatctg 180  
 46 acgttcgacc tgggtgcatcg cctatggctc gacgccggac acaccgtgca gtacgtccag 240  
 48 aacgtcaccg acgtggacga cccgttggtc gagcgtgctg agcgcgacgg catcgactgg 300  
 50 cggacgctgg gcgaccgcga gacgcagctg ttccgtgagg acatggccgc gttgcgcgtg 360  
 52 ctgccccgcg acgactacgt cgccgcgacc gacgcgatcg ccgaggctcg cgagatggtc 420  
 54 gagaagctgc tggcctcggg tgcggcgtag atcgtcgagg acgccgagta ccccgacgtg 480  
 56 tactttccgc cgcgacccac cgcgcagttc gggtacgagt ccggctacga ccgcgacacc 540  
 58 atgctcacgt tgttcgccga acgcggcggg gaccgcgacc gcccgggcaa gtccgatcaa 600  
 60 ctgcagcgtg tgctgtggcg cgccgagcgt cctggcgagc ccagctggcc ttccgcgttc 660  
 62 ggccggggcc ggcccggctg gcacgtggaa tgttcggcga tcgccctgac gcggatcggc 720  
 64 accggcctcg acatccaggg cggcggcagc gacctcatct tcccgcacca cgagtattcg 780  
 66 gccgcgcacg ccgaatccgt caccggtgag cgacgattcg cagccacta cgtgcacacc 840  
 68 ggcatgatcg gctgggacgg ccacaagatg agcaagagcc gcggcaacct ggtcctgggtg 900  
 70 tcgcagttgc gcgccaggg cgtcgaccgg tcggcgatcc ggctcggcct gttctccggg 960

## RAW SEQUENCE LISTING

DATE: 02/02/2006

PATENT APPLICATION: US/10/511,244

TIME: 12:34:04

Input Set : A:\UCSD1420-1.ST25.txt

Output Set: N:\CRF4\02022006\J511244.raw

```

72 cactaccgcg aggaccgggt ctggagcaac gaggttctcg acgaggccaa cgcgcgactc 1020
74 gcgcggtggc gcagtgccac cgcattgccc gaggcgccc atgcgaccga cgtgatcgcg 1080
76 cgcgtccggc agtacctggc cgatgacctg gacacgccga aagcgcttgc cgcactcgat 1140
78 ggttggtgta ccgacgcgct gtctacgggt gggcacgaca ccgagtcgcc gcggctcgtg 1200
80 gccaccaccg tcgacgcggt gctgggtgtg gacctc 1236
83 <210> SEQ ID NO: 2
84 <211> LENGTH: 412
85 <212> TYPE: PRT
86 <213> ORGANISM: Mycobacterium smegmatis
88 <400> SEQUENCE: 2
90 Met Gln Ser Trp Ser Ala Pro Ala Ile Pro Val Val Pro Gly Arg Gly
91 1 5 10 15
94 Pro Ala Leu Arg Leu Phe Asp Ser Ala Asp Arg Gln Val Arg Pro Val
95 20 25 30
98 Thr Pro Gly Pro Thr Ala Thr Met Tyr Val Cys Gly Ile Thr Pro Tyr
99 35 40 45
102 Asp Ala Thr His Leu Gly His Ala Ala Thr Tyr Leu Thr Phe Asp Leu
103 50 55 60
106 Val His Arg Leu Trp Leu Asp Ala Gly His Thr Val Gln Tyr Val Gln
107 65 70 75 80
110 Asn Val Thr Asp Val Asp Asp Pro Leu Phe Glu Arg Ala Glu Arg Asp
111 85 90 95
114 Gly Ile Asp Trp Arg Thr Leu Gly Asp Arg Glu Thr Gln Leu Phe Arg
115 100 105 110
118 Glu Asp Met Ala Ala Leu Arg Val Leu Pro Pro His Asp Tyr Val Ala
119 115 120 125
122 Ala Thr Asp Ala Ile Ala Glu Val Val Glu Met Val Glu Lys Leu Leu
123 130 135 140
126 Ala Ser Gly Ala Ala Tyr Ile Val Glu Asp Ala Glu Tyr Pro Asp Val
127 145 150 155 160
130 Tyr Phe Arg Ala Asp Ala Thr Ala Gln Phe Gly Tyr Glu Ser Gly Tyr
131 165 170 175
134 Asp Arg Asp Thr Met Leu Thr Leu Phe Ala Glu Arg Gly Gly Asp Pro
135 180 185 190
138 Asp Arg Pro Gly Lys Ser Asp Gln Leu Asp Ala Leu Leu Trp Arg Ala
139 195 200 205
142 Glu Arg Pro Gly Glu Pro Ser Trp Pro Ser Pro Phe Gly Arg Gly Arg
143 210 215 220
146 Pro Gly Trp His Val Glu Cys Ser Ala Ile Ala Leu Thr Arg Ile Gly
147 225 230 235 240
150 Thr Gly Leu Asp Ile Gln Gly Gly Gly Ser Asp Leu Ile Phe Pro His
151 245 250 255
154 His Glu Tyr Ser Ala Ala His Ala Glu Ser Val Thr Gly Glu Arg Arg
155 260 265 270
158 Phe Ala Arg His Tyr Val His Thr Gly Met Ile Gly Trp Asp Gly His
159 275 280 285
162 Lys Met Ser Lys Ser Arg Gly Asn Leu Val Leu Val Ser Gln Leu Arg
163 290 295 300
166 Ala Gln Gly Val Asp Pro Ser Ala Ile Arg Leu Gly Leu Phe Ser Gly

```

## RAW SEQUENCE LISTING

DATE: 02/02/2006

PATENT APPLICATION: US/10/511,244

TIME: 12:34:04

Input Set : A:\UCSD1420-1.ST25.txt

Output Set: N:\CRF4\02022006\J511244.raw

```

167 305          310          315          320
170 His Tyr Arg Glu Asp Arg Phe Trp Ser Asn Glu Val Leu Asp Glu Ala
171          325          330          335
174 Asn Ala Arg Leu Ala Arg Trp Arg Ser Ala Thr Ala Leu Pro Glu Ala
175          340          345          350
178 Pro Asp Ala Thr Asp Val Ile Ala Arg Val Arg Gln Tyr Leu Ala Asp
179          355          360          365
182 Asp Leu Asp Thr Pro Lys Ala Leu Ala Ala Leu Asp Gly Trp Cys Thr
183          370          375          380
186 Asp Ala Leu Ser Tyr Gly Gly His Asp Thr Glu Ser Pro Arg Leu Val
187 385          390          395          400
190 Ala Thr Thr Val Asp Ala Leu Leu Gly Val Asp Leu
191          405          410
194 <210> SEQ ID NO: 3
195 <211> LENGTH: 1242
196 <212> TYPE: DNA
197 <213> ORGANISM: Mycobacterium tuberculosis
199 <400> SEQUENCE: 3
200 atgcagtcgt ggtattgccc accggttcg gtgttgccgg gacgaggccc gcagctacgg      60
202 ctgtacgaca gcgcgaccg gcaggtcggt ccggtggcgc ccgatctaa ggccaccatg      120
204 tacgtctgcg ggatcacgcc ctacgacgcc acgcatctgg gccatgctgc cacctatgtg      180
206 acgttcgacc tgatccatcg gctgtggctg gatctcgggtc atgaattgca ctatgtccag      240
208 aacatcaccg acatcgacga tccactatgt gagcgcgcgg atcgcgacgg tgtcgactgg      300
210 cgtgaccttg cccaagccga ggtcgccctg ttctgtgagg acatggcggc gctgcggggtg      360
212 ctaccaccgc aagactacgt gggggccacc gaagcgattg ctgaaatggg cgagctcatc      420
214 gaaaaaatgc tggcgtgcgg ggcggcctat gtcatagacc gggaaatggg agagtaccag      480
216 gacatctact tccgcgctga cgccaccctg cagttcgggt acgagtcagg gtatgaccgt      540
218 gacaccatgc tgcggctgtg cgaggaacgt ggcggcgac cgcggcgccc cggcaagagc      600
220 gacgaactcg acgcgttggt gtggcgggcc gcgcggcccg gtgagcccag ctggccgtcc      660
222 ccgttcgggc ctggccggcc aggctggcat gtcgagtgcg cagccatcgc gctcagtcgt      720
224 atcgggaagcg gcctcgacat ccagggcggg ggtagcgatc tgatctttcc gcaccacgag      780
226 ttcaccgctg cgcacgccga atgtgtcagc ggcaacggc gattcgcgcg gcactacgtg      840
228 catgccggga tgatcggctg ggacgggcac aagatgtcaa agagccgcgg caacctcgtg      900
230 ctggtgtcgg cgctgcgtgc gcaggacgtt gagccatcgg cggttcgggt gggtttgctc      960
232 gccggacact accgagccga tcggttctgg agccagcagg tgcttgacga ggcgaccgcc      1020
234 cggtgcacc gttggcgcac cgcaaccgca cttcccgcgc gtccggccgc agttgacgtt      1080
236 gtcgctcggg tgcgcgcta cctggccgac gatctcgata cgcctaaagc gattgccgca      1140
238 ctggatggtt gggtcaccga tgcggtggag tacggcggcc acgatgccgg ggcgcgaag      1200
240 ttggtggcga cggcgatcga tgccctgctc ggggtggacc tg      1242
243 <210> SEQ ID NO: 4
244 <211> LENGTH: 414
245 <212> TYPE: PRT
246 <213> ORGANISM: Mycobacterium tuberculosis
248 <400> SEQUENCE: 4
250 Met Gln Ser Trp Tyr Cys Pro Pro Val Pro Val Leu Pro Gly Arg Gly
251 1          5          10          15
254 Pro Gln Leu Arg Leu Tyr Asp Ser Ala Asp Arg Gln Val Arg Pro Val
255          20          25          30
258 Ala Pro Gly Ser Lys Ala Thr Met Tyr Val Cys Gly Ile Thr Pro Tyr

```

## RAW SEQUENCE LISTING

DATE: 02/02/2006

PATENT APPLICATION: US/10/511,244

TIME: 12:34:04

Input Set : A:\UCSD1420-1.ST25.txt

Output Set: N:\CRF4\02022006\J511244.raw

```

259          35          40          45
262 Asp Ala Thr His Leu Gly His Ala Ala Thr Tyr Val Thr Phe Asp Leu
263          50          55          60
266 Ile His Arg Leu Trp Leu Asp Leu Gly His Glu Leu His Tyr Val Gln
267 65          70          75          80
270 Asn Ile Thr Asp Ile Asp Asp Pro Leu Phe Glu Arg Ala Asp Arg Asp
271          85          90          95
274 Gly Val Asp Trp Arg Asp Leu Ala Gln Ala Glu Val Ala Leu Phe Cys
275          100          105          110
278 Glu Asp Met Ala Ala Leu Arg Val Leu Pro Pro Gln Asp Tyr Val Gly
279          115          120          125
282 Ala Thr Glu Ala Ile Ala Glu Met Val Glu Leu Ile Glu Lys Met Leu
283          130          135          140
286 Ala Cys Gly Ala Ala Tyr Val Ile Asp Arg Glu Met Gly Glu Tyr Gln
287 145          150          155          160
290 Asp Ile Tyr Phe Arg Ala Asp Ala Thr Leu Gln Phe Gly Tyr Glu Ser
291          165          170          175
294 Gly Tyr Asp Arg Asp Thr Met Leu Arg Leu Cys Glu Glu Arg Gly Gly
295          180          185          190
298 Asp Pro Arg Arg Pro Gly Lys Ser Asp Glu Leu Asp Ala Leu Leu Trp
299          195          200          205
302 Arg Ala Ala Arg Pro Gly Glu Pro Ser Trp Pro Ser Pro Phe Gly Pro
303          210          215          220
306 Gly Arg Pro Gly Trp His Val Glu Cys Ala Ala Ile Ala Leu Ser Arg
307 225          230          235          240
310 Ile Gly Ser Gly Leu Asp Ile Gln Gly Gly Gly Ser Asp Leu Ile Phe
311          245          250          255
314 Pro His His Glu Phe Thr Ala Ala His Ala Glu Cys Val Ser Gly Glu
315          260          265          270
318 Arg Arg Phe Ala Arg His Tyr Val His Ala Gly Met Ile Gly Trp Asp
319          275          280          285
322 Gly His Lys Met Ser Lys Ser Arg Gly Asn Leu Val Leu Val Ser Ala
323          290          295          300
326 Leu Arg Ala Gln Asp Val Glu Pro Ser Ala Val Arg Leu Gly Leu Leu
327 305          310          315          320
330 Ala Gly His Tyr Arg Ala Asp Arg Phe Trp Ser Gln Gln Val Leu Asp
331          325          330          335
334 Glu Ala Thr Ala Arg Leu His Arg Trp Arg Thr Ala Thr Ala Leu Pro
335          340          345          350
338 Ala Gly Pro Ala Ala Val Asp Val Val Ala Arg Val Arg Arg Tyr Leu
339          355          360          365
342 Ala Asp Asp Leu Asp Thr Pro Lys Ala Ile Ala Ala Leu Asp Gly Trp
343          370          375          380
346 Val Thr Asp Ala Val Glu Tyr Gly Gly His Asp Ala Gly Ala Pro Lys
347 385          390          395          400
350 Leu Val Ala Thr Ala Ile Asp Ala Leu Leu Gly Val Asp Leu
351          405          410
354 <210> SEQ ID NO: 5
355 <211> LENGTH: 404

```

## RAW SEQUENCE LISTING

DATE: 02/02/2006

PATENT APPLICATION: US/10/511,244

TIME: 12:34:04

Input Set : A:\UCSD1420-1.ST25.txt

Output Set: N:\CRF4\02022006\J511244.raw

```

356 <212> TYPE: PRT
357 <213> ORGANISM: Corynebacterium striatum
359 <400> SEQUENCE: 5
361 Met His Ala Trp Pro Asp Pro Ser Val Pro Ala Val Ala Gly Thr Pro
362 1 5 10 15
365 Val Pro Leu Lys Leu Phe Asp Thr Ala Asp Gln Arg Val Lys Glu Val
366 20 25 30
369 Asp Thr Thr Pro Asp Ala Asn Gly Glu Val Gly Met Tyr Val Cys Gly
370 35 40 45
373 Ile Thr Pro Tyr Asp Ser Thr His Leu Gly His Ala Ala Thr Tyr Leu
374 50 55 60
377 Thr Phe Asp Leu Ala Gln Arg Gln Leu Leu Ala Asn Gly His Lys Val
378 65 70 75 80
381 His Tyr Val Gln Asn Ile Thr Asp Val Asp Asp Pro Leu Phe Glu Arg
382 85 90 95
385 Ala Glu Arg Asp Gly Val Asp Trp Arg Glu Leu Gly Thr Ser Gln Ile
386 100 105 110
389 Asn Leu Phe Arg Ser Asp Met Glu Ile Leu Ser Val Ile Pro Pro Cys
390 115 120 125
393 Asp Tyr Ile Gly Ala Met Glu Ser Val Asp Glu Val Ile Ala Met Val
394 130 135 140
397 Gln Gln Leu Leu Asp Ala Gly Ala Ala Tyr Glu Leu Asp Gln Gly Asp
398 145 150 155 160
401 Ile Tyr Ala Ser Ile Asp Ala Thr Glu Gln Phe Gly Tyr Glu Ser Asn
402 165 170 175
405 Leu Asp Arg Ala Thr Met Glu Glu Tyr Phe Ala Glu Arg Gly Gly Asp
406 180 185 190
409 Pro Asp Arg Glu Gly Lys Arg Asp Pro Leu Asp Ala Leu Val Trp Arg
410 195 200 205
413 Gly His Arg Glu Gly Glu Pro Ala Trp Asp Ser Pro Phe Gly Pro Gly
414 210 215 220
417 Arg Pro Gly Trp His Val Glu Cys Ser Ala Ile Ala Thr Asn Arg Leu
418 225 230 235 240
421 Gly Ser His Phe Ala Ile Gln Gly Gly Gly Ser Asp Leu Ala Phe Pro
422 245 250 255
425 His His Glu Phe Ser Ala Ala His Ala Glu Ala Ala Leu Lys Val Glu
426 260 265 270
429 Arg Met Ala Gly His Tyr Val His Ala Gly Met Ile Ala Leu Asp Gly
430 275 280 285
433 Val Lys Met Ser Lys Ser Leu Gly Asn Leu Val Phe Val His Lys Leu
434 290 295 300
437 Ser Glu Ala Gly His Asp Pro Ser Ala Ile Arg Leu Ala Val Phe Ala
438 305 310 315 320
441 Gly His Tyr Arg Glu Asp Arg Asp Phe Ser Asp Ala Ile Leu Ala Glu
442 325 330 335
445 Ala Glu Glu Arg Leu Thr Arg Trp Arg Glu Gln Leu Ala Gly Glu Val
446 340 345 350
449 Ser Glu Ala Glu Ala Thr Glu Val Asp Lys Leu Arg Ala Ile Leu
450 355 360 365

```

RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 02/02/2006

PATENT APPLICATION: US/10/511,244

TIME: 12:34:05

Input Set : A:\UCSD1420-1.ST25.txt

Output Set: N:\CRF4\02022006\J511244.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:8; Xaa Pos. 1,2,13

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/511,244

DATE: 02/02/2006

TIME: 12:34:05

Input Set : A:\UCSD1420-1.ST25.txt

Output Set: N:\CRF4\02022006\J511244.raw

L:19 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:616 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0

L:1843 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:26